CHAPTER 3

USING PROBABILISTIC ANALYSIS IN HUMAN HEALTH ASSESSMENT

3.0 INTRODUCTION

This chapter outlines how probabilistic analysis may be applied to human health risk assessments in the Environmental Protection Agency's (EPA) Superfund program. The paradigm for human health risk assessment as described in EPA's *Risk Assessment Guidance for Superfund* (U.S. EPA, 1989), includes data collection/evaluation in addition to exposure and toxicity assessment and risk characterization. Although the strategies and methods used in collecting and analyzing data can significantly impact the uncertainty in a risk estimate, they are issues relevant to risk assessment in general, and are addressed in other guidance documents, such as EPA's *Guidance for Data Useability in Risk Assessment* (U.S. EPA, 1992b). RAGS Volume 3: Part A focuses on a tiered approach for incorporating quantitative information on variability and uncertainty into risk management decisions.

3.1 CHARACTERIZING VARIABILITY IN EXPOSURE VARIABLES

Exhibit 3-1 gives the general equation used for calculating exposure, often expressed as an average daily intake. In a point estimate approach, single values (typically a mixture of average and high-end values) are input into the equation. In probabilistic risk assessment (PRA), the only difference is that a probability distribution, rather than single value, is specified for one or more variables. A Monte Carlo simulation is executed by repeatedly selecting random values from each of these distributions and calculating the corresponding exposure and risk. For the majority of PRAs, it is expected that probability distributions will be used to characterize inter-individual variability, which refers to true heterogeneity or diversity in a population. Thus, variability in daily intake, for example, can be characterized by combining multiple sources of variability in exposure, such as ingestion rate, exposure frequency, exposure duration, and body weight. Variability in chemical concentrations (Chapter 5 and Appendix C) and the toxicity term in ecological risk assessment (Chapter 4) may also be considered in risk calculations.

EXHIBIT 3-1
GENERAL EQUATION FOR EXPOSURE

$$I = \frac{C \times CR \times EF \times ED}{BW \times AT} \qquad \text{Eq. 3-1}$$

where,

I	=	daily intake
C	=	contaminant concentration
CR	=	contact rate (ingestion, inhalation, dermal contact)
EF	=	exposure frequency
ED	=	exposure duration
BW	=	body weight
AT	=	averaging time

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3

95% UCL for mean - The one-sided 95% upper confidence limit for a population mean; if a sample of size (*n*) was repeatedly drawn from the population, the 95% UCL will equal or exceed the true population mean 95% of the time. It is a measure of uncertainty in the mean, not to be confused with the 95th percentile (see below), which is a measure of variability. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged.

95th percentile - The number in a distribution that is greater than 95% of the other values of the distribution, and less than 5% of the values. When estimated from a sample, this quantity may be equal to an observed value, or interpolated from among two values.

Arithmetic Mean (AM) - A number equal to the average value of a population or sample. Usually obtained by summing all the values in the sample and dividing by the number of values (i.e., sample size).

Assessment Endpoint - The specific expression of the population or ecosystem that is to be protected. It can be characterized both qualitatively and quantitatively in the risk assessment.

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in the population, usually considered to be the arithmetic mean or median of the risk distribution.

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

CTE Risk - The estimated risk corresponding to the central tendency exposure.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF or PMF, gives the cumulative probability of occurrence for a random independent variable. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*.

Exposure Point Concentration (EPC) - The average chemical concentration to which receptors are exposed within an exposure unit. Estimates of the EPC represent the concentration term used in exposure assessment.

Frequency Distribution/Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

High-end Risk - A risk descriptor representing the high-end, or upper tail of the risk distribution, usually considered to be equal to or greater than the 90th percentile.

Low-end Risk - A risk descriptor representing the low-end, or lower tail of the risk distribution, such as the 5th or 25th percentile.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Probability Density Function (PDF) - A function representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Probability Mass Function (PMF) - A function representing the probability distribution for a discrete random variable. The mass at a point refers to the probability that the variable will have a value at that point.

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3—*Continued*

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic *r* that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity.*
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic *r* that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Target Population - The set of all receptors that are potentially at risk. Sometimes referred to as the "population of concern". A sample population is selected for statistical sampling in order to make inferences regarding the target population (see Appendix B, Section B.3.1, Concepts of Populations and Sampling).

Figure 3-1 shows a hypothetical example of an input distribution for drinking water ingestion rate. Assume that survey data for drinking water ingestion rates were compiled in order to select and fit a probability distribution. One of the first steps in exploring the data set may be to plot a frequency distribution. In the graph, the height of the bars (the y-axis) represents the relative frequency of ingestion rates in the population and the spread of the bars (the x-axis) is the varying amounts of water ingested (L/day). Since ingestion rate is a continuous random variable, the probability distribution can also be represented graphically with a probability density function (PDF). Assume that the following parameters are estimated from the sample: arithmetic mean=1.36, standard deviation=0.36, geometric mean=1.31, and geometric standard deviation=1.30. These parameter estimates may be used to define a variety of probability distributions, including a 2-parameter lognormal distribution. The fit of the lognormal distribution can be evaluated by visual inspection using the PDF given by Figure 3-1, or by a lognormal probability plot (see Appendix B).

The y-axis for a PDF is referred to as the *probability density*, where the density at a point on the x-axis represents the probability that a variable will have a value within a narrow range about the point. This type of graph shows, for example, that there is a greater area under the curve (greater probability density) in the 1-2 L/day range than 0-1 L/day or 2-3 L/day. That is, most people reported consuming 1-2 L/day of drinking water. By selecting a lognormal distribution to characterize inter-individual variability, we can state more precisely that 1 L/day corresponds to the 15th percentile and 2 L/day corresponds to the 95th percentile, so approximately 80% (i.e., 0.95–0.15=0.80) of the population is likely to consume between 1 and 2 L/day of drinking water.

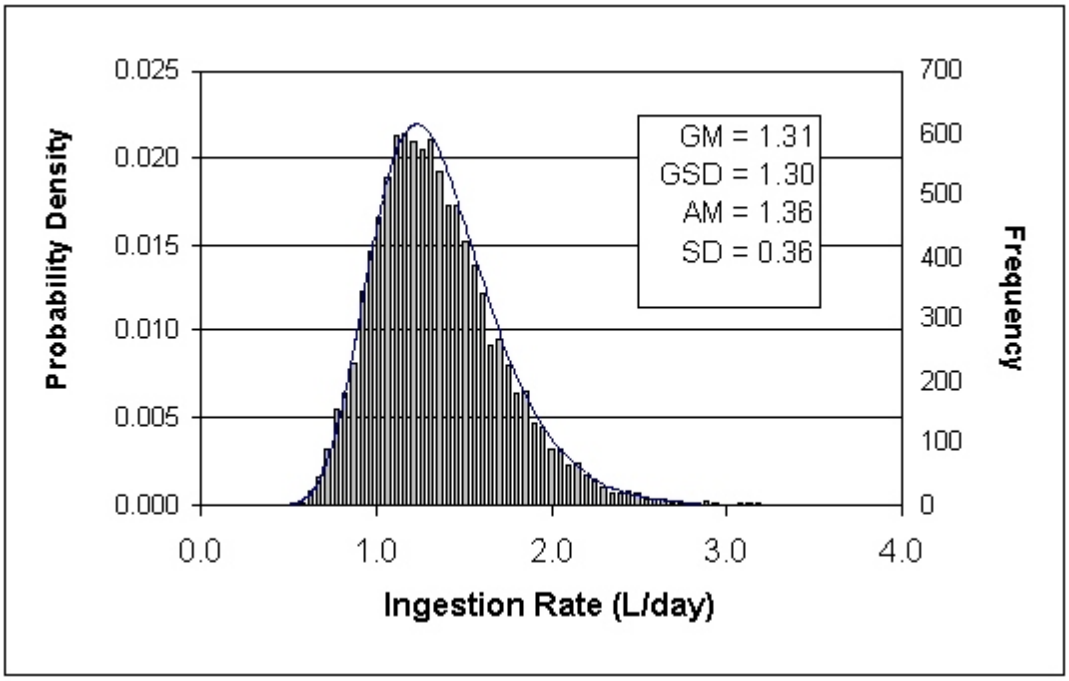


Figure 3-1. Example of a frequency distribution for adult drinking water ingestion rates, overlaid by a graph of the probability density function (PDF) for a lognormal distribution defined by the sample statistics. The distribution represents inter-individual variability in water intakes and is characterized by two parameters. Typically, the geometric mean (GM) and geometric standard deviation (GSD), or the arithmetic mean (AM) and arithmetic standard deviation (SD) are presented to characterize a lognormal distribution.

3.1.1 DEVELOPING DISTRIBUTIONS FOR EXPOSURE VARIABLES

When site-specific data or representative surrogate data are available, a probability distribution can be fit to that data to characterize variability. Appendix B describes how to fit distributions to data, how to assess the quality of the fit and discusses topics such as the sensitivity of the tails of the distribution to various PDFs, and correlations among variables. Many of the issues discussed below regarding the use of site-specific data or surrogate data are relevant to both point estimate risk assessment and PRA.

For the majority of the exposure variables, such as exposure duration, water intake rates, and body weight, site-specific data will not be available. The risk assessor will have to either select a distribution from existing sources, or develop a distribution from published data sets and data summaries. Examples of sources for these distributions and data sets are EPA's *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), Oregon Department of Environmental Quality's *Guidance for Use of Probabilistic Analysis in Human Health Risk Assessment* (Oregon DEQ, 1998), and the scientific literature. An appropriate PDF should be determined in collaboration with the regional risk assessor. The process by which PDFs are to be selected and evaluated should be described in the workplan. EPA's Superfund program is in the process of developing a ranking methodology to evaluate data representativeness relevant to various exposures scenarios. Following peer review and project completion, the results will be posted on EPA Superfund web page.

☞ *At this time, EPA does not recommend generic or default probability distributions for exposure variables.*

Regardless of whether a PDF is derived from site-specific measurements or obtained from the open literature, the risk assessor should carefully evaluate the applicability of the distribution to the target population at the site. The distribution selected should be derived from the target population or from a surrogate population that is representative of the target population at the site. For example, a distribution based on homegrown vegetable consumption in an urban population would not be representative for a farming population in the Midwest. If such a distribution were to be used, (and no other data were available), the uncertainty and bias that this PDF would impart to the risk estimate should be communicated to the risk decision makers.

For purposes of risk management decision making, the significance of not having site-specific data should be evaluated in the context of representativeness and sensitivity analysis. If published data are representative of the potentially exposed population, then site-specific data may be unnecessary. For example, body weights of children and adults have been well studied from national surveys and can generally be considered reasonable surrogates for use in site risk assessments. Furthermore, even if a variable is likely to vary among different exposed populations, it may not contribute greatly to the variance or uncertainty in risk estimates. In this case, surrogate data may also be used with confidence in the risk estimate. In addition, the PRA may be simplified by using point estimates instead of probability distributions for the "less sensitive" exposure variables. In part, the decision to use a point estimate in lieu of a probability distribution must balance the benefit of simplifying the analysis and the communication process (see Chapter 6), against the reduction (however small) in the variance of the risk distribution. The utility of sensitivity analysis in identifying the important factors in a risk estimate is discussed further below and in Appendix A.

It is also important to evaluate the sample design and sample size when deciding to apply a distribution to a specific site. Depending on the situation, a very large data set derived from a national population may be more useful than a site-specific data set derived from a small, incomplete, or poorly designed study. Appendix B provides additional discussion on how to evaluate the data and studies that form the basis for a distribution. Often, the question arises regarding the appropriateness of combining data sets to derive a PDF. Before combining data sets, a careful evaluation should be made of the representativeness of the study populations, and the similarity in study designs and quality. In addition, statistical tests may be used to determine whether or not data sets are compatible with a common probability distribution (Hedges and Olkin, 1985; Stiteler et al., 1993). In general, risk assessors should be reluctant to combine data sets for the purpose of developing a PDF that characterizes variability. Due to the number of potential differences inherent in the study design, alternative data sets may provide a better measure of uncertainty in the probability distribution and parameter estimates, rather than a means of increasing the overall sample size for defining a single probability distribution. For example, if multiple data sets are available, a more informative approach may be to incorporate each data set into the PRA in a separate analysis, as a form of sensitivity analysis on the choice of alternative data sets.

Each probability distribution used in a Monte Carlo Analysis (MCA) should be presented with sufficient detail that the analysis can be reproduced (see Chapter 1, Section 1.4, Condition #2). This information may be presented in tabular and/or graphical summaries. Important information for a summary table would include a description of the distribution type (e.g., lognormal, gamma, etc.), the parameters that define the distribution (e.g., mean and standard deviation, and possibly upper and lower truncation limits for a normal distribution), units, and appropriate references (see Table 3-6, for example). The table should also indicate whether the distribution describes variability or uncertainty. The report should discuss the representativeness of the data and why a particular data set was selected if alternatives were available. Graphical summaries of the distributions may include both PDFs and cumulative distribution functions (CDFs), and should generally be used to document distributions that characterize site-specific data.

3.1.2 CHARACTERIZING RISK USING PRA

Quantitative risk characterization involves evaluating exposure (or intake) estimates against a benchmark of toxicity, such as a cancer slope factor or a noncancer hazard quotient. The general equation used for quantifying cancer risk from ingestion of contaminated soil is shown in Exhibit 3-3, and the equation for noncarcinogenic hazard is shown in Exhibit 3-4. A Hazard Index is equal to the sum of chemical-specific Hazard Quotients.

At this time, this guidance does not propose probabilistic approaches for dose-response in human health assessment and, further, *discourages undertaking such activities on a site-by-site basis*. Such activities require contaminant-specific national consensus development and national policy development (see Chapter 1, Section 1.4.1). Chapter 4 discusses methods for applying probabilistic approaches to ecological dose-response assessment.

The probabilistic calculation of risk involves random sampling from each of the exposure variable distributions. The output of this process is a distribution of risk estimates. When the calculation of risk (or any other model endpoint) is repeated many times using Monte Carlo techniques to sample the variables at random, the resulting distribution of risk estimates can be displayed in a similar fashion. The type of summary graph used to convey the results of a MCA depends on the risk management needs. For example, Chapter 1, Figure 1-3 shows how a PDF for risk might be used to compare the probabilistic

estimate of the RME risk (e.g., 95th percentile) with a risk level of concern. This type of summary can also be used to effectively illustrate the relationship between the RME risk determined from point estimate and probabilistic approaches.

EXHIBIT 3-3

EQUATION FOR CANCER RISK

$$Risk = Dose \times CSF$$

Example for Soil Ingestion

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

where,

C	=	concentration in soil (mg/kg)	ED	=	exposure duration (years)
IR	=	soil ingestion rate (mg/day)	BW	=	body weight (kg)
CF	=	conversion factor (1E-06 kg/mg)	AT	=	averaging time (days)
EF	=	exposure frequency (days/year)	CSF	=	oral cancer slope factor $(mg/kg\text{-}day)^{-1}$

EXHIBIT 3-4

EQUATION FOR NONCANCER HAZARD QUOTIENT

$$Hazard\ Quotient = \frac{Dose}{RfD}\ or\ \frac{Concentration}{RfC}$$

where,

RfD	=	reference dose, oral or dermally adjusted (mg/kg-day)
RfC	=	reference concentration, inhalation (g/m^3)

In addition, the CDF can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., cancer risk of 1E-04 or Hazard Index of 1). Figure 3-2 illustrates both the PDF and CDF for risk for a hypothetical scenario. Factors to consider when applying the PDF or CDF are discussed in Chapter 1, Exhibit 1-3. When in doubt about the appropriate type of summary to use, both the PDF and CDF should be provided for all risk distributions. At a minimum, each summary output for risk should highlight the risk descriptors of concern (e.g., 50th, 90th, 95th, and 99.9th percentiles). It can also be informative to include the results of the point estimate analysis—the risks corresponding to the central tendency exposure (CTE) and the reasonable maximum exposure (RME).

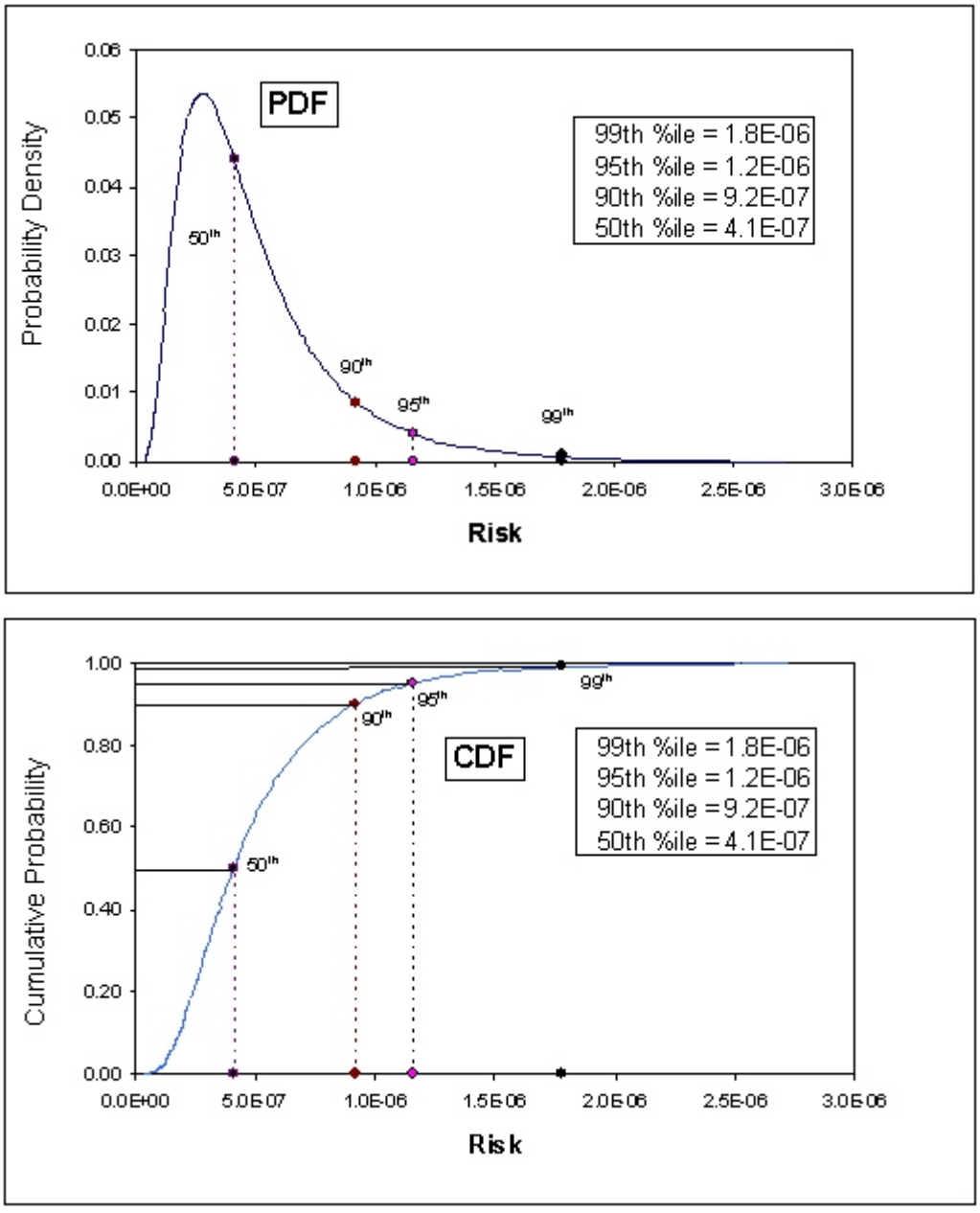


Figure 3-2. Hypothetical PRA results showing a PDF (top panel) and CDF (bottom panel) for cancer risk with selected summary statistics. The CDF rises to a maximum cumulative probability of 1.0. The CDF clearly shows that the level of regulatory concern chosen for this example (1E-06) falls between the 90th and 95th percentiles of the risk distribution.

3.2 ROLE OF THE SENSITIVITY ANALYSIS

Prior to conducting a PRA, it is worthwhile to review several points pertaining to the sensitivity analysis. As shown in Chapter 2 (Figures 2-1 and 2-2), sensitivity analysis can play an important role in decision making at each tier of the tiered process. Beginning with Tier 1, a point estimate for risk should be calculated prior to conducting a PRA. Based on the results of the point estimate, the risk assessor and risk decision makers should determine whether a probabilistic analysis will offer additional benefit. One factor in this decision may be the results of a sensitivity analysis. A primary objective of the sensitivity analysis is to determine which variables and pathways most strongly influence the risk estimate. At many Superfund sites, an estimate of cumulative risk considers contamination in multiple media, moving through multiple pathways and interacting with a number of receptors. Depending on the complexity of the site, and the modeling approaches, a risk assessment may involve one exposure pathway and few variables, or multiple pathways with many variables (e.g., multimedia fate and transport models). However, resources and time are often limited. The sensitivity analysis is invaluable in focusing these limited resources on the most influential variables and pathways.

Several methods for conducting sensitivity analysis are described in Appendix A. It is important to note that when a sensitivity analysis is performed and the major variables are identified, this does not mean that the less influential pathways and variables should be eliminated from the risk assessment. It means that because they are not major contributors to the variability or uncertainty in risk, they can be described with point estimates without affecting the risk management decision. If distributions are readily available for these less influential variables, one may use distributions. The key goal is to provide a comprehensive risk characterization that is scientifically credible and sufficient for risk decision making. The time and effort required to achieve various levels of complexity should be weighed against the value of the information provided to the risk managers.

Additionally, if a variable is specified as influential in the sensitivity analysis, this does not automatically mean that a distribution has to be developed for this variable. If the risk assessor feels that data are simply not sufficient from which to develop a distribution, then a plausible point estimate can be used. The risk assessor should be aware of a possible problem arising from using point estimates in the absence of data adequate to support a distribution. If a variable has the potential to significantly impact the risk outcome, and a very high-end or low-end point estimate is used in the PRA, this has the potential to right-shift or left-shift the final distribution of risk. Even though there might not be enough data to develop a distribution of variability for an influential variable, it would be prudent to communicate the importance of this data gap to the risk decision makers, and perhaps run multiple simulations with several plausible input distributions for that variable. Communication of this uncertainty may persuade the risk decision makers to collect additional data to better define the variable.

3.3 EXPOSURE POINT CONCENTRATION TERM

A brief discussion of the concentration term is provided below. A more complete discussion of the concentration term in PRA is provided in Appendix C. The reader is also referred to Chapter 5 on development of PRGs.

The major source of uncertainty in Superfund risk assessments is often incomplete knowledge of the concentration of one or more chemicals in various exposure media. In any risk assessment, the derivation of the concentration term will reflect assumptions about: (1) properties of the contaminant, (2) the spatial and temporal variability in contamination, (3) the behavior of the receptor, and (4) the time scale of the toxicity of the chemical(s).

Contaminant concentrations contacted by a receptor are likely to vary depending on the spatial variability of contamination and the movements of the receptor. Different individuals may be exposed to different concentrations based on inter-individual variability in activity patterns. If information regarding activity patterns is unavailable, receptors are typically assumed to exhibit random movement such that there is an equal probability of contacting any area within the exposure unit (EU). An EU is defined as the geographical area in which a receptor moves and contacts contaminated medium during the period of the exposure duration. In addition, in Superfund risk assessments, the toxicity criteria are often based on health effects associated with chronic exposure (e.g., lifetime risk of cancer following chronic daily intake over a period of 30 years). Hence, the most appropriate expression for the concentration term, for the majority of risk assessments, is one that characterizes the long-term average exposure point concentration within the EU.

> ☞ *The most appropriate expression of the exposure point concentration term for chronic exposure will characterize the long-term average concentration experienced by a receptor within the exposure unit.*

In point estimate risk assessments, the exposure point concentration term is usually calculated as the 95% upper confidence limit (95% UCL) of the arithmetic mean because of the uncertainty associated with estimating the true (i.e., population) mean concentration at a site. If the sampling density is sparse relative to the size of the EU, the uncertainty may be high due to the relatively small number of measurements available to estimate the mean concentration within the EU. The decision to use the upper confidence limit to define the concentration term introduces a measure of protectiveness by reducing the chance of underestimating the mean. Although there will be situations in which modeling variability in concentration will be the appropriate choice (e.g., non-random movement within an EU, acute exposure events, migration of groundwater contaminant plume, migration of fish, etc.), in most cases, characterization of the concentration term will focus on uncertainty. Appendix C provides a more complete discussion on characterizing both variability and uncertainty in the concentration term. Table 3-1 summarizes a number of appropriate methods for characterizing uncertainty in the parameter of an exposure variable, such as the arithmetic mean of the concentration term.

3.4 CHARACTERIZING UNCERTAINTY IN EXPOSURE VARIABLES

Uncertainty is described as a lack of knowledge about factors affecting exposure or risk. To evaluate regulatory options, risk assessors are expected to translate the available evidence, however tentative, into a probability of occurrence of an adverse health effect. Data from a sample or surrogate population are used to develop estimates of exposure and risk in a specific target population (see Section 3.1.4 and Appendix B, Section B.3.1). This extrapolation requires assumptions and inferences that have inherent strengths and limitations, and may bias the outcome of the risk estimate. For example, a common assumption in risk assessments for carcinogens is that a contaminant concentration within the boundaries of a hazardous waste site represents the concentration that a receptor is exposed to throughout the period of exposure, with the corresponding dose averaged over the course of a lifetime. This assumption may be conservative (i.e., result in overestimation of exposure) if it is unlikely that receptors will be exposed at the hazardous waste site for the entire exposure duration. It is incumbent on the risk assessor to clearly present the rationale for the assumptions used in a risk assessment, as well as their implications and limitations.

U.S. EPA guidance, including the *Exposure Assessment Guidelines* (U.S. EPA, 1992a), *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), and *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997d) have classified uncertainty in exposure assessment into three broad categories:

(1) *Parameter uncertainty* - uncertainty in values used to estimate variables of a model;
(2) *Model uncertainty* - uncertainty about a model structure (e.g., exposure equation) or intended use; and
(3) *Scenario uncertainty* - uncertainty regarding missing or incomplete information to fully define exposure.

Each source of uncertainty is described in detail below, along with strategies for addressing them in PRA.

3.4.1 PARAMETER UNCERTAINTY

Parameter uncertainty may be the most readily recognized source of uncertainty that is quantified in site-specific risk assessments at hazardous waste sites. Parameter uncertainty can occur in each step of the risk assessment process from data collection and evaluation, to the assessment of exposure and toxicity. Sources of parameter uncertainty may include systematic errors or bias in the data collection process, imprecision in the analytical measurements, and extrapolation from surrogate measures to represent the parameter of interest. For example, soil data collected only from the areas of highest contamination, rather than the entire area that a receptor is expected to come into contact, will result in a biased estimate of exposure.

In general, parameter uncertainty can be quantified at any stage of the tiered process, including point estimate analysis (Tier 1), one-dimensional Monte Carlo analysis (1-D MCA) (Tier 2), and two-dimensional Monte Carlo analysis (2-D MCA) (Tier 3). In the point estimate approach, parameter uncertainty may be addressed in a qualitative manner for most variables. For example, the uncertainty section of a point estimate risk assessment document might state that an absorption fraction of 100% was used to represent the amount of contaminant in soil absorbed from the gastrointestinal (GI) tract, and as a result, the risk estimate may overestimate actual risk. In addition, a sensitivity analysis may be performed, wherein one input variable at a time is changed, while leaving the others constant, to examine

the effect on the outcome. In the case of absorption from the GI tract, different plausible estimates of the high-end, or RME absorption fraction might be used as inputs to the risk equation. The differences in the risk estimates would reflect uncertainty in the RME absorption fraction.

Quantitative approaches for characterizing parameter uncertainty in exposure variables in a Monte Carlo simulation are summarized in Table 3-1. If uncertainty in only a few parameter values is of interest, multiple 1-D MCA simulations can yield the same results as a 2-D MCA simulation, but without the time and effort of a 2-D MCA. An example illustrating this concept is given in Table 3-2. With multiple 1-D MCA simulations, variability is characterized in one or more variables using probability distributions for variability (PDFv's), and uncertainty in a parameter is characterized with a series of different point estimates from a probability distribution for uncertainty (PDFu) (e.g., 95% lower confidence limit LCL [95% LCL], sample mean, and 95% UCL). In a 2-D MCA simulation, variability is characterized in one or more variables using PDFv's, and uncertainty in one or more parameters is characterized with PDFu's. With both approaches, the influence of the parameter uncertainty can be presented as a credible interval or confidence interval (CI) around the risk distribution, depending on how the PDFu's are defined. When only a few sources of parameter uncertainty are quantified, multiple 1-D MCA simulations are preferred over a 2-D MCA because the approach is easier to use and communicate. However, if the goal is to explore the effect that many sources of parameter uncertainty may have on the risk estimates simultaneously, a 2-D MCA is preferred. Iterative 1-D MCA simulations with different combinations of confidence limits may be impractical.

Table 3-1. Methods for Characterizing Parameter Uncertainty with Monte Carlo Simulations.

Approach	Example of Model Input	Method	Example of Model Output
Single Point Estimate	• 95% UCL	1-D MCA	PDFv[1] for risk, calculated using the 95% UCL for one parameter.
Multiple Point Estimates	• 95% LCL • sample mean • 95% UCL	1-D MCA	Three PDFv's for risk, representing the 90% CI for each percentile of the risk distribution.[2] The 90% CI only accounts for uncertainty in a single parameter (not multiple parameters).
Parametric PDFu[1]	PDFu for the mean based on the sampling distribution, derived from a Student's *t*-distribution.	2-D MCA	One PDFv for risk with confidence intervals at each percentile of the risk distribution. The CI reflects uncertainty in one or more parameters.
Non-parametric PDFu	PDFu for the mean based on bootstrap resampling methods.	2-D MCA	Same as parametric probability distribution for uncertainty.

[1]Probability distribution for uncertainty (PDFu) and probability distribution for variability (PDFv).
[2]The 95% UCL for the concentration term represents a 1-sided confidence interval (CI), meaning there is a 95% probability that the value is *greater* than or equal to the mean. Similarly, the 95% LCL would represent the 1-sided CI in which there is a 95% probability that the value is *less* than or equal to the mean. Both values are percentiles on the probability distribution for uncertainty (PDFu), also called the sampling distribution for the mean. Together, the 95% LCL and 95% UCL are equal to the 2-sided 90% confidence interval only for cases in which the PDFu is symmetric. For example, the sampling distribution for the arithmetic mean of a sample from a normal distribution with an unknown variance is described with the symmetric Student's *t*-distribution, whereas the PDFu for the mean of a lognormal distribution is asymmetric. In order to compare the results of multiple 1-D MCA simulations and a 2-D MCA simulation, the same methodology should be employed to define the PDFu and the corresponding confidence limits.

It is generally incorrect to combine a PDFu for one parameter (e.g., mean of the concentration term) with one or more PDFv's in other exposure factors when conducting a 1-D MCA for variability. However, distributions for uncertainty and variability may be appropriately combined in a 2-D MCA. As discussed in Appendix D, with 2-D MCA, a clear distinction should be made between probability distributions that characterize variability (PDFv) and parameter uncertainty (PDFu). A 2-D MCA propagates the uncertainty and variability distributions separately through an exposure model, thereby making it possible to evaluate the effect of each on the risk estimates.

Example: Comparison of Multiple Point Estimates of Uncertainty in 1-D MCA, and Distributions of Uncertainty in 2-D MCA

Table 3-2 illustrates an application of the approaches presented in Table 3-1 for quantifying variability and parameter uncertainty. This is a hypothetical example, and no attempt was made to use standard default assumptions for exposure variables. Two sources of variability are quantified: (1) inter-individual variability in exposure frequency (EF), characterized by a triangular distribution, and (2) inter-individual variability in exposure duration (ED), characterized by a truncated lognormal distribution. In addition, two sources of uncertainty are presented: (1) a point estimate for soil and dust ingestion rate, intended to characterize the RME; and (2) an upper truncation limit of the lognormal distribution for ED, intended to represent a plausible upper bound for the exposed population. Methods for quantifying these sources of uncertainty are discussed below. Additional sources of uncertainty may also have been explored. For example, the choice of a triangular distribution for a PDFv may be provocative for some risk assessors, since there are few cases in which empirical data suggest a random sample is from a triangular distribution. Nevertheless, triangular distributions may be considered rough, or "preliminary" distributions (see Chapter 2 and Appendix B, Section B.2) for cases when the available information supports a plausible range and central tendency.

The choice of distributions is a potential source of uncertainty that can be explored by rerunning simulations with each alternative, plausible choice, and examining the effect on the RME risk. Simulations with preliminary simulations may yield at least three different outcomes. First, this type of sensitivity analysis can help guide efforts to improve characterizations of variability for selected variables that have the greatest affect on the risk estimates. Second, results may provide justification to exit the tiered process without continuing with additional Monte Carlo simulations since further effort would be unlikely to change the risk management decision. Finally, if the major sources of uncertainty can be clearly identified, a subset of the less sensitive variables may be defined by point estimates without significantly reducing the uncertainty in the risk estimates.

Parameter uncertainty can be quantified for both point estimates and PDFv's. In this example, both types of inputs (i.e., point estimates and PDFv's) are presented as sources of parameter uncertainty: the RME point estimate for soil and dust ingestion rate (IRsd), and the upper truncation limit on a PDFv for ED. For IRsd, assume that three different studies provide equally plausible values for the RME: 50, 100, and 200 mg/day. A uniform PDFu is specified to characterize this range of plausible values. For ED, assume that the maximum value reported from a site-specific survey was 26 years, but surrogate data for other populations suggest the maximum may be as long as 40 years. A uniform PDFu is specified to characterize this range of plausible values as well.

In Cases 1-3, the impact of uncertainty in IRsd and ED was evaluated using a series 1-D MCA simulations. Inputs for uncertain parameters associated with IRsd and ED in Case 1, 2, and 3 represent the minimum, central tendency, and maximum values, respectively. Each simulation yields a different

risk distribution based on different combinations of point estimates for parameters. Although a PDFu was specified for IRsd, it would have been incorrect to combine the PDFu with the PDFv's for EF and ED in a 1-D MCA because the result would have been a single distribution of risk that co-mingled uncertainty and variability.

In Case 4, a single 2-D MCA simulation was run using the PDFu's for uncertainty and the PDFv's for variability. By propagating variability and uncertainty separately, the 2-D MCA yields a series of distributions of risk, from which credible intervals can be calculated for each percentile of the CDF.

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

Table 3-2. Example of 1-D MCA and 2-D MCA.

Variable	Type of Input	1-D MCA			2-D MCA
		Case 1	Case 2	Case 3	Case 4
C (mg/kg)	pt estimate	500	500	500	500
IRsd (mg/day)	pt estimate	50	100	200	see below
	PDFu for pt estimate	--	--	--	uniform (50, 200)[a]
CF (kg/mg)	pt estimate	1E-06	1E-06	1E-06	1E-06
EF (days/year)	PDFv	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350
ED (years)	PDFv	T-lognormal mean = 9 stdv = 10 max = 26	T-lognormal mean = 9 stdv = 10 max = 33	T-lognormal mean = 9 stdv = 10 max = 40	T-lognormal mean = 9 stdev = 10 max = PDFu (see below)
	PDFu for parameter of PDFv	--	--	--	max ~ uniform (26, 40)[b]
BW (kg)	pt estimate	70	70	70	70
AT (days)	pt estimate	25550	25550	25550	25550
CSF $(mg/kg\text{-}day)^{-1}$	pt estimate	1E-01	1E-01	1E-01	1E-01

[a]Uncertainty in the RME point estimate, defined by a uniform distribution with parameters (minimum, maximum).
[b]Uncertainty in the upper truncation limit of the lognormal distribution, defined by a PDFv with parameters (mean, standard deviation, maximum) and a PDFu for the maximum defined by a uniform distribution with parameters (minimum, maximum).

Monte Carlo Simulation Results

Figures 3-3 and 3-4 illustrate CDFs for risk produced from Monte Carlo simulations using *Crystal Ball*® 2000. The 1-D MCA simulations (Figure 3-3) were run with 10,000 iterations and Latin Hypercube sampling. The 2-D MCA simulation (Figure 3-4) was run with 250 iterations of the outer loop (uncertainty) and 2,000 iterations of the inner loop (variability). Details regarding 2-D MCA simulation are given in Appendix D.

Figure 3-3 shows CDFs for risk based on three simulations of a 1-D MCA simulation. Each simulation used a different combination of plausible estimates of the RME value for IRsd and the upper truncation limit for ED, as discussed above. The results provide a bounding estimate on the risk distribution given these two sources of uncertainty. The 95th percentile risk, highlighted as an example of the RME risk estimate, may range from approximately 7E-06 to 3.5E-05.

Figure 3-4 shows a single CDF for risk, representing the central tendency risk distribution. This CDF was derived by simulating uncertainty in the risk distribution using 2-D MCA. For this example, the 2-D MCA yields 250 simulations of the risk distributions for variability, so that there are 250 plausible estimates of each percentile of the risk distribution. In practice, more than 250 simulations may be needed to adequately quantify uncertainty in the risk distribution. Results of a 2-D MCA can be presented as probability distributions of uncertainty, or box-and-whisker plots of uncertainty at selected percentiles of the risk distributions. Figure 3-4 shows the central tendency (50th percentile) estimate of uncertainty for the entire CDF of risk. In addition, a box-and-whisker plot is shown at the 95th percentile of the CDF. Selected statistics for the box-and-whisker plot are included in a text box on the graphic (i.e., minimum; 5th, 50th, and 95th percentiles, and maximum). The 90% credible interval is given by the 5th and 95th percentiles. For this example, the 90% credible interval for the 95th percentile of the risk distribution is: [9.1E-06, 3.1E-05].

Figures 3-3 and 3-4 demonstrate that the two approaches (i.e., multiple 1-D MCA and 2-D MCA) can yield the same results. However, when there are numerous sources of uncertainty, 2-D MCA offers at least two advantages over multiple 1-D MCA simulations: (1) 2-D MCA allows the multiple sources of uncertainty to be included simultaneously so the approach is more efficient than a series of 1-D MCA simulations; and (2) multiple 1-D MCA simulations yield multiple estimates of the RME risk, but it is not possible to characterize the uncertainty in the RME risk in quantitative terms; a 2-D MCA yields a PDFu for RME risk, which allows for statements regarding the level of certainty that the RME risk is above or below a risk level of concern.

The 95th percentile is a focus of this example because it is a recommended starting point for determining the risk corresponding to the RME. Chapter 7 provides guidance to the risk decision makers on choosing an appropriate percentile (on a distribution of variability) within the RME risk range (90th to 99.9th percentiles). The chapter also includes a qualitative consideration of the uncertainty or confidence surrounding a risk estimate in the decision-making process.

Figure 3-3

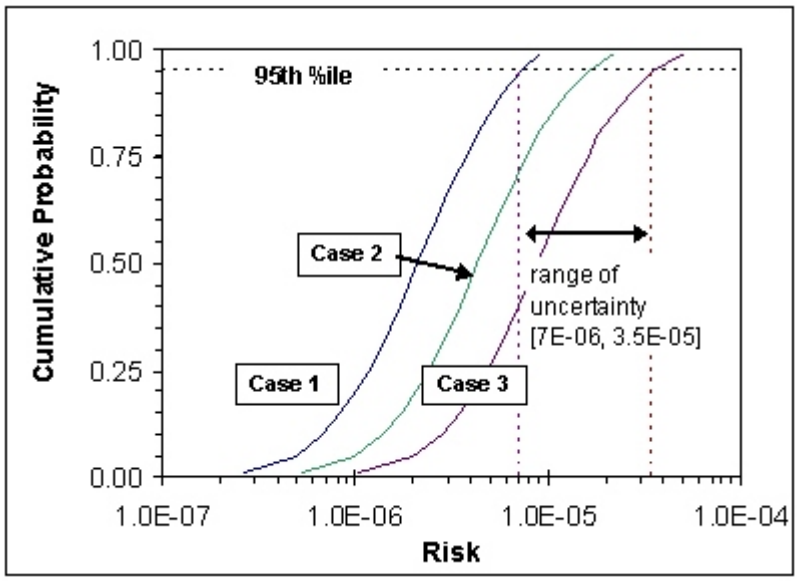
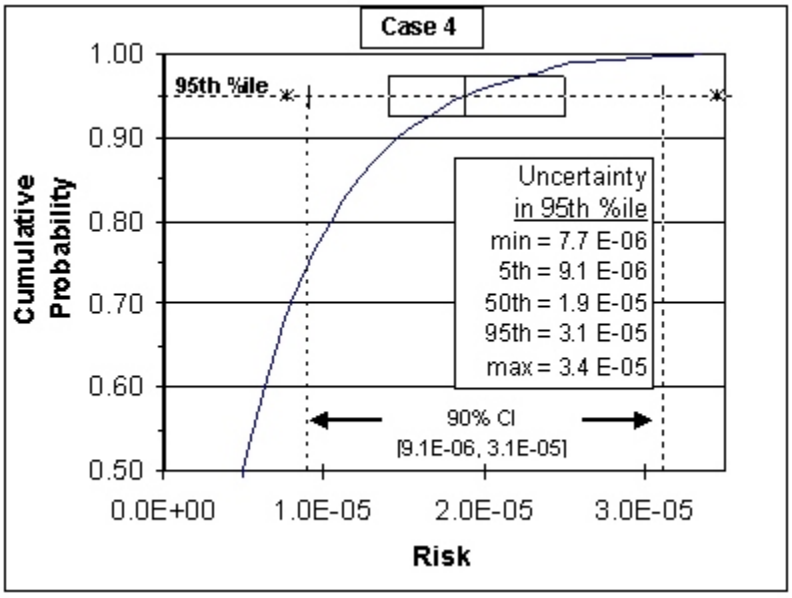


Figure 3-4



3.4.2 SCENARIO AND MODEL UNCERTAINTY

All models are simplified representations of complex biological and physical processes. As such, they, and the scenarios to which they are applied, may introduce a significant source of uncertainty into an exposure and risk estimate. Models may exclude important variables or important pathways of exposure, ignore interactions between inputs, use surrogate variables that are different from the target variables, or they may be designed for specific scenarios and not others. As a result, a model may not adequately represent all aspects of the phenomena it was intended to approximate or it may not be appropriate to predict outcomes for a different type of scenario. For example, a model intended to estimate risk from continuous, steady state exposures to a contaminant may not be appropriate or applicable for estimating risk from acute or subchronic exposure events. In any risk assessment, it is important to understand the original intent of a model, the assumptions being made in a model, what the parameters represent, and how they interact. Based on this knowledge, one can begin to understand how representative and applicable (or inapplicable) a model may be to a given scenario. If multiple models exist that can be applied to a given scenario, it may be useful to compare and contrast results in order to understand the potential implications of the differences. The use of multiple models, or models with varying levels of sophistication, may provide valuable information on the uncertainty introduced into a risk estimate as the result of model or scenario uncertainty. The collection of measured data as a reality check against a given parameter or the predicted model outcome (such as the collection of vegetable and fruit contaminant data to compare against modeled uptake into plants) is also useful in attempting to reduce or at least gain a better understanding of model and scenario uncertainty.

3.5 EXAMPLE OF PRA FOR HUMAN HEALTH

The following hypothetical example provides a conceptual walk-through of the tiered approach for PRA in Superfund risk assessment. The example begins with a baseline human health point estimate risk assessment (Tier 1) and moves to Tier 2, in which multiple iterations of a 1-D MCA are run using default and site-specific assumptions for input distributions. The general concepts associated with the tiered approach are discussed in Chapter 2, and a similar example for ecological risk assessment is given in Chapter 4. The 1-D MCA results are based on simulations with *Crystal Ball*® 2000 using 10,000 iterations and Latin Hypercube sampling. These settings were sufficient to obtain stability (i.e., <1% difference) in the 95% percentile risk estimate. The example is presented in Exhibit 3-5. Tables and figures supporting the example are given immediately following the exhibit.

EXHIBIT 3-5
USING THE TIERED PROCESS FOR PRA
HYPOTHETICAL CASE STUDY FOR HUMAN HEALTH RISK ASSESSMENT

RI Planning/Scoping/Problem Formulation/Data Collection

- Site Description: Former federal facility
- Site Size: 100 acres (5 acres within spill area (ISA); 95 acres outside spill area (OSA))
- Stakeholders: Refuge employees, environmental activists, etc.
- Land Use: Future wildlife refuge
- Receptors: Future wildlife refuge workers (i.e., ornithologists and fishery biologists)
- Sampling Data: n=35 surface soil samples (see Figure 3-5 for sample locations)
- Chemical of Concern: ChemX
- Chemical Properties: Nonvolatile
- Toxicological Properties: Carcinogen: CSF_{oral} and CSF_{dermal} = 5.5E-02, CSF_{inh} = 2.73E-02; Noncarcinogenic health data are lacking
- Risk Level of Concern: 1E-04 for cancer

↓

Tier 1 Point Estimate - **Baseline Risk Assessment**

- Exposure Unit: (see Figure 3-5) ornithologist (exposed in OSA) and fishery biologist (exposed in ISA)
- Exposure Pathways: Ingestion of soil/dust; inhalation of fugitive dust, dermal absorption
- Concentration Term: 95% UCL for arithmetic mean (Table 3-3)
- Risk Equations: Exhibit 3-6
- Exposure Parameters: Table 3-4
- Results: Table 3-5

↓

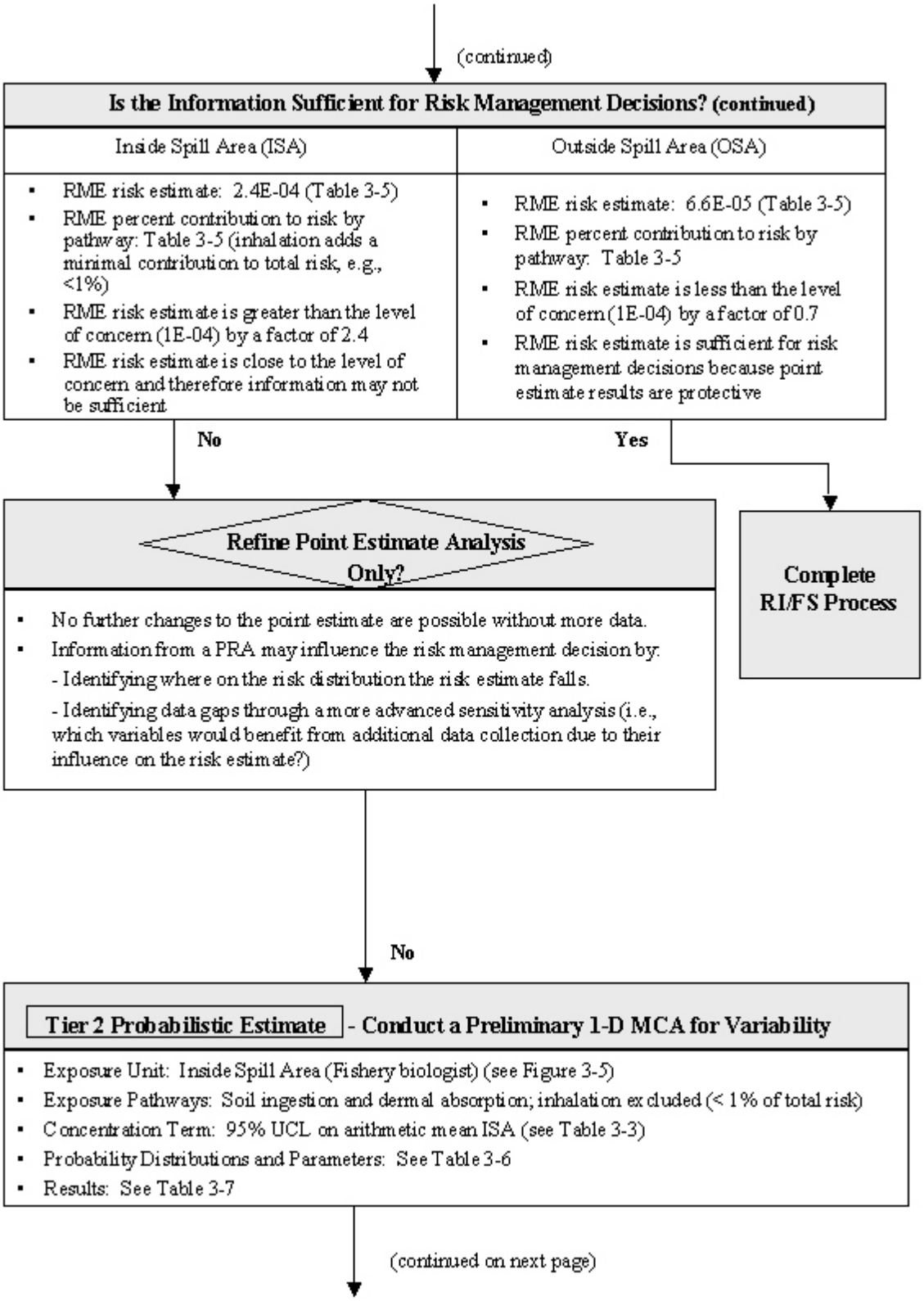
SMDP Is the Information Sufficient for Risk Management Decisions?

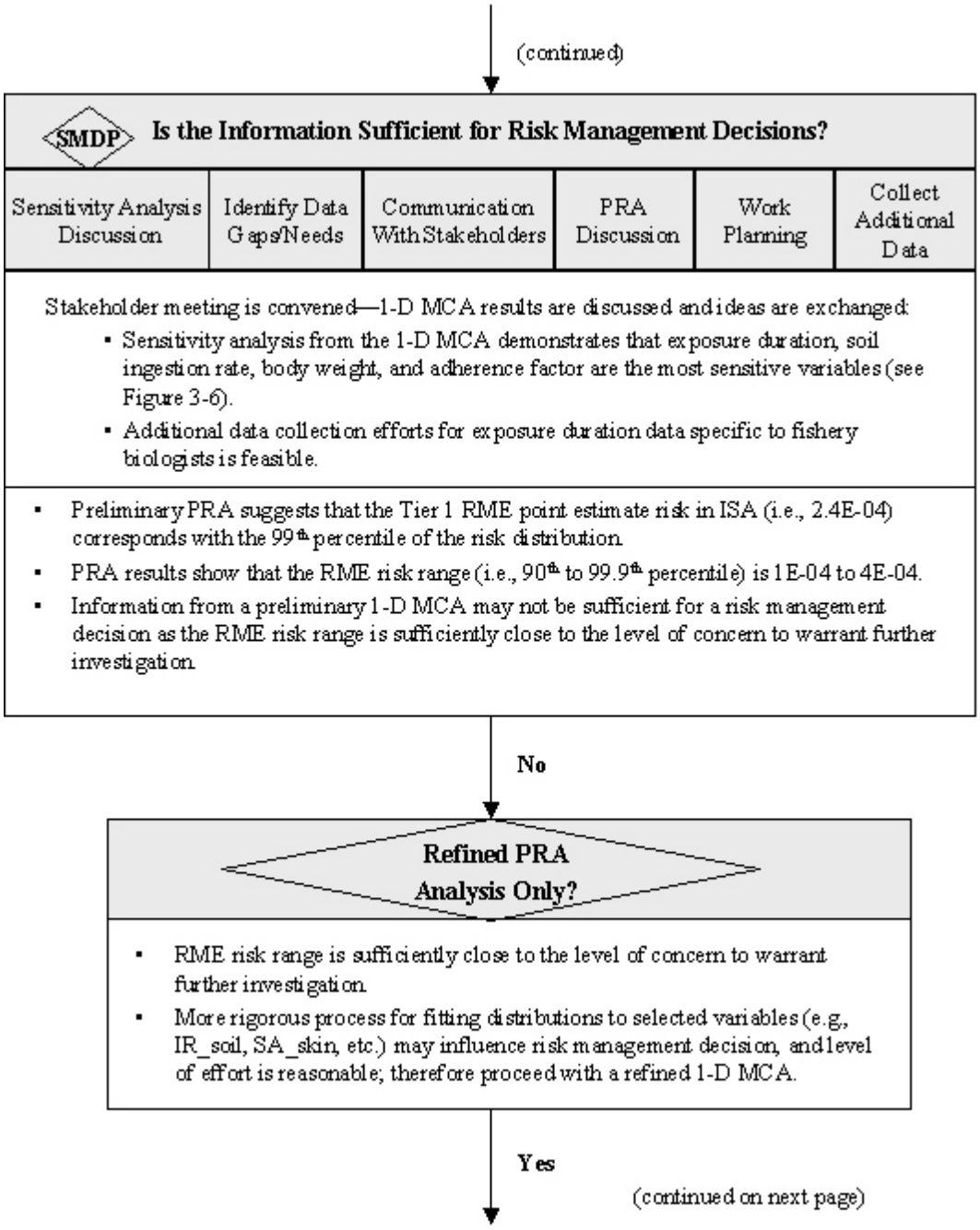
Sensitivity Analysis Discussion	Identify Data Gaps/Needs	Communication With Stakeholders	PRA Discussion	Work Planning	Collect Additional Data

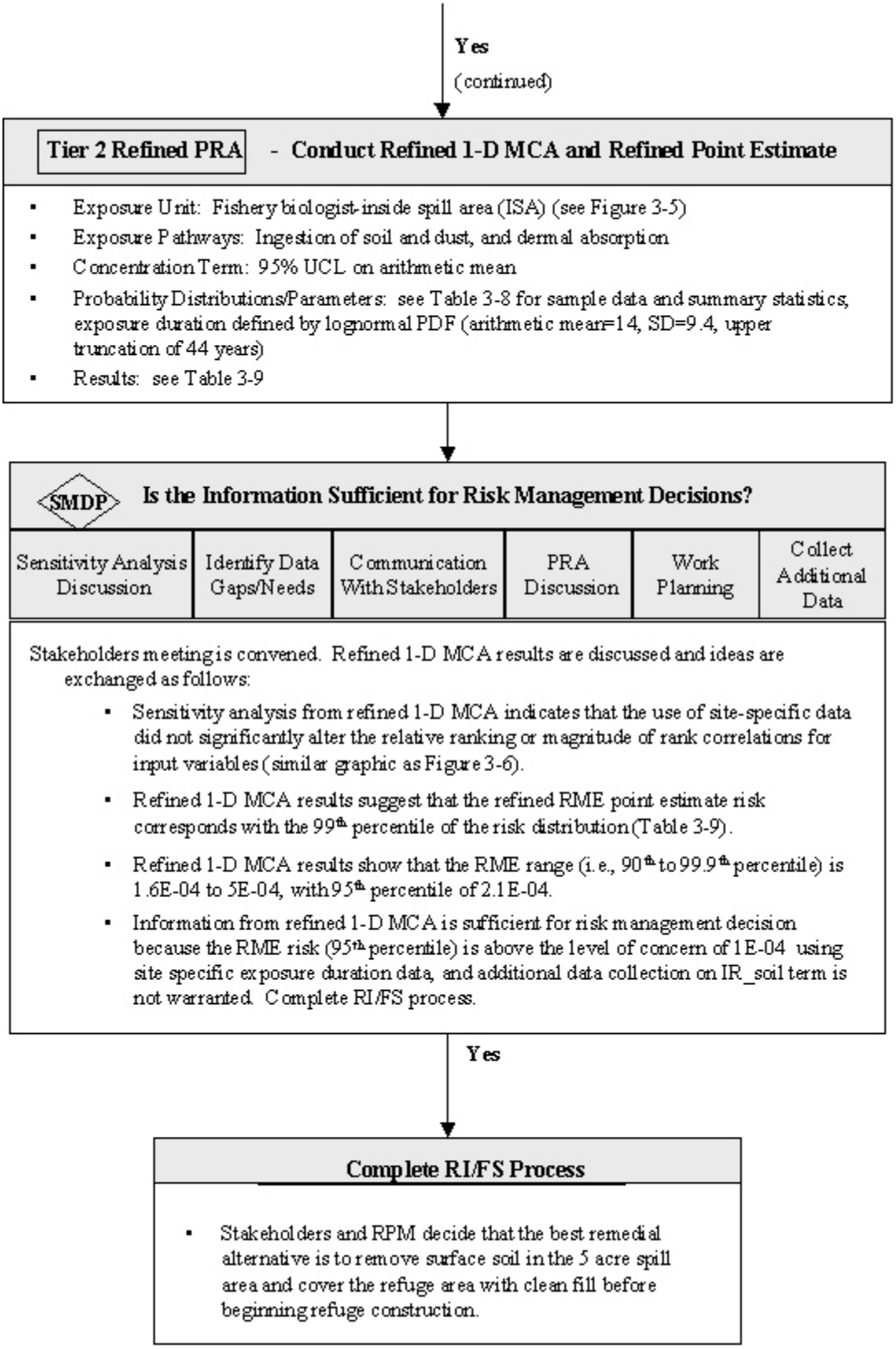
Stakeholder meeting is convened—point estimate results are discussed and ideas are exchanged as follows:

- Risk estimates are expected to be conservative due to the use of standard default exposure parameters, but are the defaults representative?
- Stakeholders are concerned about risk to workers and about the consequences of remediation (e.g., negative impacts on habitat and potential job losses).
- Stakeholders are concerned about the relevance of some nonsite-specific exposure variables (e.g., exposure duration), but are not sure which variables to investigate further (i.e., which is the most influential?).
- Results of the sensitivity analysis from point estimate risk assessment cannot identify where the high end risk estimate falls on the risk distribution.
- There is sufficient information (e.g., arithmetic mean, standard deviation, percentiles) for some of the exposure variables to develop initial probability distributions to characterize variability.

(continued on next page)


(continued)
Is the Information Sufficient for Risk Management Decisions? (continued)
Inside Spill Area (ISA)
RME risk estimate: 2.4E-04 (Table 3-5)
RME percent contribution to risk by pathway: Table 3-5 (inhalation adds a minimal contribution to total risk, e.g., <1%)
RME risk estimate is greater than the level of concern (1E-04) by a factor of 2.4
RME risk estimate is close to the level of concern and therefore information may not be sufficient
Outside Spill Area (OSA)
RME risk estimate: 6.6E-05 (Table 3-5)
RME percent contribution to risk by pathway: Table 3-5
RME risk estimate is less than the level of concern (1E-04) by a factor of 0.7
RME risk estimate is sufficient for risk management decisions because point estimate results are protective
No
Yes
Refine Point Estimate Analysis Only?
No further changes to the point estimate are possible without more data.
Information from a PRA may influence the risk management decision by:
- Identifying where on the risk distribution the risk estimate falls.
- Identifying data gaps through a more advanced sensitivity analysis (i.e., which variables would benefit from additional data collection due to their influence on the risk estimate?)
Complete RI/FS Process
No
Tier 2 Probabilistic Estimate - Conduct a Preliminary 1-D MCA for Variability
Exposure Unit: Inside Spill Area (Fishery biologist) (see Figure 3-5)
Exposure Pathways: Soil ingestion and dermal absorption; inhalation excluded (< 1% of total risk)
Concentration Term: 95% UCL on arithmetic mean ISA (see Table 3-3)
Probability Distributions and Parameters: See Table 3-6
Results: See Table 3-7
(continued on next page)


(continued)
SMDP
Is the Information Sufficient for Risk Management Decisions?
Sensitivity Analysis Discussion
Identify Data Gaps/Needs
Communication With Stakeholders
PRA Discussion
Work Planning
Collect Additional Data
Stakeholder meeting is convened—1-D MCA results are discussed and ideas are exchanged:
• Sensitivity analysis from the 1-D MCA demonstrates that exposure duration, soil ingestion rate, body weight, and adherence factor are the most sensitive variables (see Figure 3-6).
• Additional data collection efforts for exposure duration data specific to fishery biologists is feasible.
• Preliminary PRA suggests that the Tier 1 RME point estimate risk in ISA (i.e., 2.4E-04) corresponds with the 99th percentile of the risk distribution.
• PRA results show that the RME risk range (i.e., 90th to 99.9th percentile) is 1E-04 to 4E-04.
• Information from a preliminary 1-D MCA may not be sufficient for a risk management decision as the RME risk range is sufficiently close to the level of concern to warrant further investigation.
No
Refined PRA Analysis Only?
• RME risk range is sufficiently close to the level of concern to warrant further investigation.
• More rigorous process for fitting distributions to selected variables (e.g., IR_soil, SA_skin, etc.) may influence risk management decision, and level of effort is reasonable; therefore proceed with a refined 1-D MCA.
Yes
(continued on next page)


Yes
(continued)
Tier 2 Refined PRA - Conduct Refined 1-D MCA and Refined Point Estimate
Exposure Unit: Fishery biologist-inside spill area (ISA) (see Figure 3-5)
Exposure Pathways: Ingestion of soil and dust, and dermal absorption
Concentration Term: 95% UCL on arithmetic mean
Probability Distributions/Parameters: see Table 3-8 for sample data and summary statistics, exposure duration defined by lognormal PDF (arithmetic mean=14, SD=9.4, upper truncation of 44 years)
Results: see Table 3-9
SMDP
Is the Information Sufficient for Risk Management Decisions?
Sensitivity Analysis Discussion
Identify Data Gaps/Needs
Communication With Stakeholders
PRA Discussion
Work Planning
Collect Additional Data
Stakeholders meeting is convened. Refined 1-D MCA results are discussed and ideas are exchanged as follows:
Sensitivity analysis from refined 1-D MCA indicates that the use of site-specific data did not significantly alter the relative ranking or magnitude of rank correlations for input variables (similar graphic as Figure 3-6).
Refined 1-D MCA results suggest that the refined RME point estimate risk corresponds with the 99th percentile of the risk distribution (Table 3-9).
Refined 1-D MCA results show that the RME range (i.e., 90th to 99.9th percentile) is 1.6E-04 to 5E-04, with 95th percentile of 2.1E-04.
Information from refined 1-D MCA is sufficient for risk management decision because the RME risk (95th percentile) is above the level of concern of 1E-04 using site specific exposure duration data, and additional data collection on IR_soil term is not warranted. Complete RI/FS process.
Yes
Complete RI/FS Process
Stakeholders and RPM decide that the best remedial alternative is to remove surface soil in the 5 acre spill area and cover the refuge area with clean fill before beginning refuge construction.

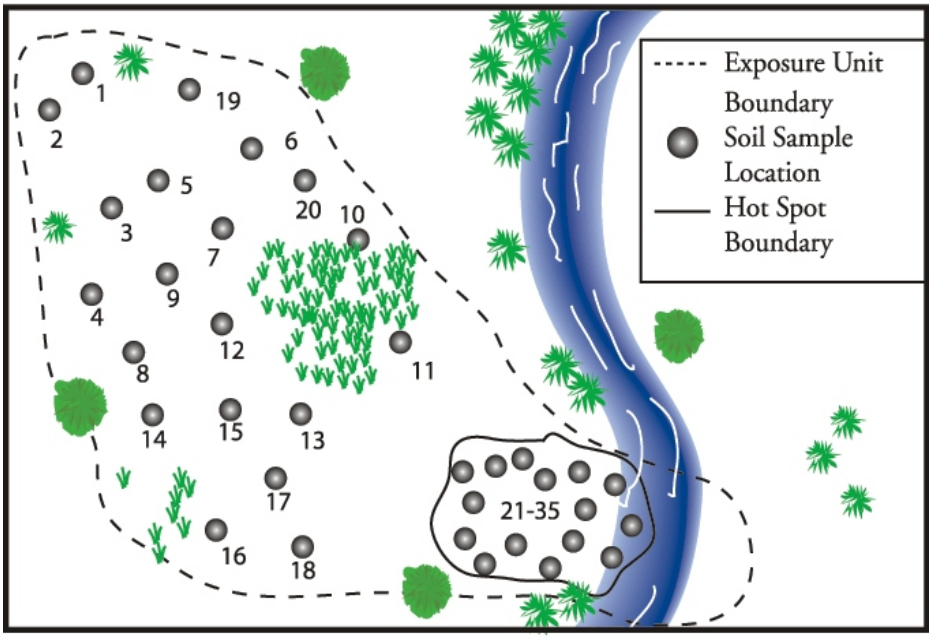


Figure 3-5. Site map for future wildlife refuge showing boundaries for the exposure unit and potential hotspot, as well as sampling locations (n=35). Sample numbers correspond with concentration data given in Table 3-3.

Table 3-3. Concentrations in Surface Soil (mg/kg).

Outside Spill Area (n=20)		Inside Spill Area (n=15)	
1088	305	1934	970
646	2787	402	985
3943	760	4215	743
149	149	1121	158
3704	1088	629	21296
845	837	2293	
488	1295	257	
387	1239	288	
1438	1006	57	
2502	283	228	

Summary Statistics	Outside Spill Area	Inside Spill Area
Mean	1247	2372
Standard Deviation	1121	5348
95% UCL[1]	2303	8444

[1]The 95% UCL was estimated using the Land method (see Appendix C).

EXHIBIT 3-6

RISK EQUATIONS

Soil Ingestion

$$\text{Risk} = \frac{Cs \times CF \times IRs \times FI \times EF \times ED}{BW \times AT} \times \text{Oral CSF}$$

Dermal Absorption

$$\text{Risk} = \frac{Cs \times CF \times SA \times AF \times ABS \times EF \times ED}{BW \times AT} \times \text{Dermal-Adjusted CSF}$$

Inhalation of Fugitive Dust

$$\text{Risk} = \frac{Cs \times 1/PEF \times IRa \times ET \times EF \times ED}{BW \times AT} \times \text{Inhalation CSF}$$

Total Risk = Sum of risks from each exposure pathway (soil + dermal + inhalation)

Where:

Cs = Concentration of ChemX in soil (mg/kg)
IRs = Soil ingestion rate for receptor (mg/day)
FI = Fraction ingested from contaminated source (unitless)
CF = Conversion factor (1E-06 kg/mg)
SA = Skin surface area available for exposure (cm^2/event)
AF = Soil to skin adherence factor for ChemX (mg/cm^2)
ABS = Absorption factor for ChemX (unitless)
IRa = Inhalation rate for receptor (m^3/hr)
PEF = Soil-to-air particulate emission factor (kg/m^3)
ET = Exposure time for receptor (hours/day)
EF = Exposure frequency for receptor (days/year)
ED = Exposure duration for receptor (years)
BW = Body weight of receptor (kg)
AT = Averaging time (years)
CSF = Cancer slope factor (oral, dermal, inhalation) $(mg/kg\text{-}day)^{-1}$

Table 3-4. Exposure Parameters used in Point Estimate Analysis.

Exposure Variable	CTE Value	RME Value	Units	Reference
IRs	50	100	mg/day	CTE: U.S. EPA, 1997a, p. 4–25 RME: U.S. EPA, 2001
FI	0.5	1	unitless	Site-specific
CF	1E-06	1E-06	kg/mg	Constant
SA	3300	3300	cm^2/event	U.S. EPA, 2001, 50th percentile value for all adult workers—exposure to face, forearms, and hands
AF	0.1	0.2	mg/cm^2	CTE: U.S. EPA, 1998; Table 3.3, value for gardeners RME: U.S. EPA, 2001
ABS	0.1	0.1	unitless	U.S. EPA, 1998, default for semi-volatile organic compounds (SVOCs)
IRa	1.3	3.3	m^3/hr	U.S. EPA, 1997a, p. 5–24, outdoor worker hourly average: mean and upper percentile
PEF	1.36E+09	1.36E+09	kg/m^3	U.S. EPA, 2001
ET	8	8	hours/day	Site-specific
EF	200	225	days/year	CTE: Site-specific assumption RME: U.S. EPA, 2001
ED	5	25	years	CTE: U.S. EPA, 1993, p. 6 RME: U.S. EPA, 2001
BW	70	70	kg	U.S. EPA, 1993, p. 7
AT	25550	25550	days	constant

CTE = central tendency exposure; RME = reasonable maximum exposure.

Table 3-5. Point Estimate Risks and Exposure Pathway Contributions.

Risk Estimate by Exposure Pathway	Inside Spill Area (n = 15)		Outside Spill Area (n = 20)	
	CTE	RME	CTE	RME
Soil Ingestion	6.5E-06 (43 %)	1.5E-04 (60 %)	1.7E-06 (43 %)	4.0E-05 (60 %)
Dermal Absorption	8.6E-06 (57 %)	9.6E-05 (40 %)	2.3E-06 (57 %)	2.6E-05 (40 %)
Inhalation	9.9E-10 (< 1 %)	1.4E-08 (< 1 %)	2.7E-10 (< 1 %)	3.8E-09 (< 1 %)
Total Risk	**1.5E-05**	**2.4E-04**	**4.1E-06**	**6.6E-05**

Example of % contribution: % Soil for RME risk inside spill area = (Soil risk / Total risk) x 100%
= (1.46E-04 / 2.42E-04) x 100% = 60%

Table 3-6. Input Distributions for Exposure Variables used in 1-D MCA for Variability.

Exposure Variable[1]	Distribution Type	Parameters[2]	Units	Reference
IR_soil	Triangular	0, 50, 100	mg/day	U.S. EPA, 1993, 2001
SA_skin[3]	Lognormal	18150, 37.4	cm^2	U.S. EPA, 1997a, Table 6-4 (Total male/female body surface area)
Absorption Fraction	Uniform	0.1, 0.2	mg/cm^2	U.S. EPA, 2001; minimum truncation limit is professional judgment
IR_air	Lognormal	1.68, 0.72	m^3/hour	U.S. EPA, 1996, p.5–10
EF	Triangular	200, 225, 250	days	U.S. EPA, 2001; truncation limits are professional judgment
ED	Lognormal[4]	11.7, 7.0	years	U.S. EPA, 1997b, Table 15-161 and U.S. EPA, 2001 (Mean value is based on average of total median tenure for professional specialty and farming, forestry, and fishing)
	Truncated Lognormal[5]	14.0, 9.4, 44.0	years	Site-specific survey data, used in refined 1-D MCA
BW	Lognormal	71.75, 14.2	kg	U.S. EPA, 1997a, Tables 7-4 and 7-5; (Combined male/female body weight distributions)

[1]All other exposure parameters are inputted as point estimates (see Table 3-4).
[2]Parameters for lognormal PDF are X ~ Lognormal (arithmetic mean, arithmetic standard deviation) unless otherwise stated. Parameters for triangular PDF are X ~ Triangular (minimum, mode, maximum). Parameters for uniform PDF are X ~ Uniform (minimum, maximum).
[3]A point estimate of 0.189 was used to adjust the surface area skin (SA_skin) distribution, which is based on total body surface area, to account for skin exposures limited to face, forearms, and hands (U.S. EPA, 1997a, Vol. I).
[4]Parameters for preliminary lognormal PDF for ED were converted from a geometric mean of 10 and a 95th percentile of 25.
[5]Parameters for site-specific lognormal PDF for ED are arithmetic mean, standard deviation, and upper truncation limit.

Table 3-7. 1-D MCA Risk Estimates using Preliminary Inputs.

Cumulative Percentile	Spill Area Risk
50th	5.7E-05
90th	1.3E-04
95th	1.6E-04
99th	2.4E-04
99.9th	3.9E-04

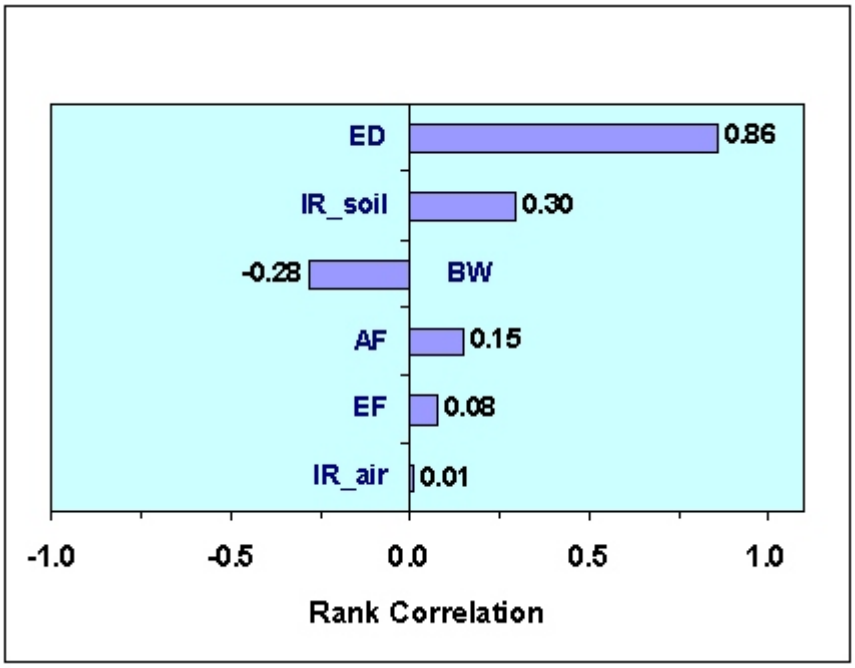


Figure 3-6. Results of sensitivity analysis for preliminary 1-D MCA (Tier 2) showing the Spearman Rank correlations (see Appendix A and B) between input variables and risk estimates.

Table 3-8. Exposure Duration Survey Results.

Survey Results (years)			Summary Statistics	
24.9	20.3	17.2	n	20
8.4	11.7	6.5	min	3.0
3.0	4.7	16.5	max	44.2
6.8	20.9	6.0	arithmetic mean	14.0
18.5	10.6	18.8	standard dev	9.4
9.1	12.7	11.7	median/GM	11.7
7.2	44.2		GSD	1.8

Table 3-9. Refined Point Estimate and 1-D MCA Risk Estimates.

Cumulative Percentile	Spill Area Risk
Refined RME Point Estimate	3.1E-04
50th	6.7E-05
90th	1.6E-04
95th	2.1E-04
99th	3.2E-04
99.9th	5.3E-04

REFERENCES FOR CHAPTER 3

Hedges, L.V. and I. Olkin. 1985. *Statistical Methods for Meta-Analysis.* Academic Press, Inc. Orlando.

Oregon DEQ. 1998. *Guidance for the Use of Probabilistic Analysis in Human Health Exposure Assessments*. Waste Management and Cleanup Division. Interim Final. November.

Stiteler, W.M., L.A. Knauf, R.C. Hertzberg, and R.S. Schoeny. 1993. A Statistical Test of Compatibility of Data Sets to a Common Dose-Response Model. *Regulatory Tox. Pharm*. 18: 392–402.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment.* EPA/600/Z-92/001. 57 *Federal Register*, 22888-22938. May 29.

U.S. EPA. 1992b. *Guidance on Data Usability in Risk Assessment*. Part A. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7.09A. NTIS PB92-96336.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund.* Office of Solid Waste and Emergency Response. Washington, DC.

U.S. EPA. 1996. *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 540/R-96/018.

U.S. EPA. 1997a. *Exposure Factors Handbook, Volume 1.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 1997b. *Exposure Factors Handbook, Volume 2.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fb.

U.S. EPA. 1997c. *Exposure Factors Handbook, Volume 3.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fc.

U.S. EPA. 1997d. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment.* Final. National Center for Environmental Assessment, Washington, DC. EPA/630/R-95/002F.

U.S. EPA. 2001. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Office of Solid Waste and Emergency Response. Washington, DC. OSWER Directive No. 9355.4-24. December.